



(11) **EP 1 229 493 A2**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
07.08.2002 Bulletin 2002/32

(51) Int Cl.7: **G06T 5/00**

(21) Application number: **01204772.6**

(22) Date of filing: **10.12.2001**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE TR
 Designated Extension States:
AL LT LV MK RO SI

(72) Inventors:
 • **Chen, Shoupu**
Rochester, New York 14650-2201 (US)
 • **Ray, Lawrence A.**
Rochester, New York 14650-2201 (US)

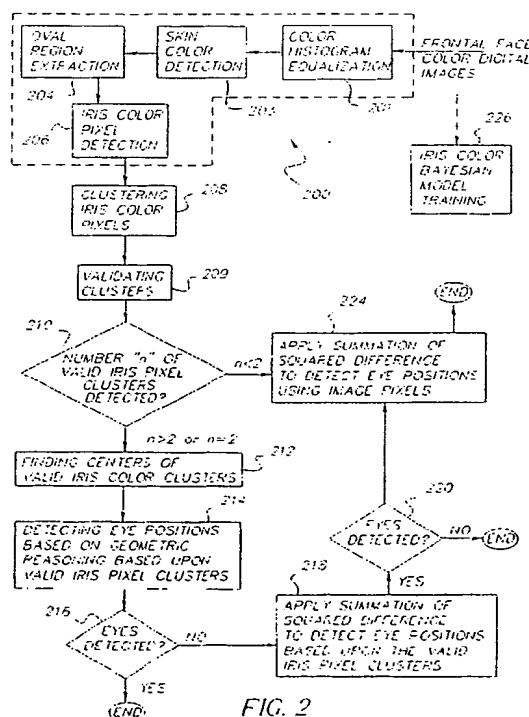
(30) Priority: **19.12.2000 US 740562**

(71) Applicant: **EASTMAN KODAK COMPANY**
Rochester, New York 14650 (US)

(74) Representative:
Nunney, Ronald Frederick Adolphe et al
Kodak Limited,
Patents, W92-3A,
Headstone Drive
Harrow, Middlesex HA1 4TY (GB)

(54) **Multi-mode digital image processing method for detecting eyes**

(57) The present invention comprises a digital image processing method for detecting human eyes in a digital image. This method comprises the steps of: detecting iris pixels in the image (200); clustering the iris pixels (208), and selecting at least one of the following methods to identify eye positions: applying geometric reasoning to detect eye positions using the iris pixel clusters (212, 214); applying a summation of squared difference method using the iris pixel clusters to detect eye positions (218); and applying a summation of squared difference method to detect eye positions from the pixels in the image (224). The method applied is selected on the basis of the number of iris pixel clusters. In another embodiment, the present invention also comprises a computer program product.



EP 1 229 493 A2

Description

[0001] The present invention relates to digital image processing methods for detecting facial features and more particularly to methods for detecting human eyes.

[0002] In digital image processing it is often useful to detect human eyes in an image. This information is used, for example, to locate other features in the image such as human visual orientation. This information can also be used for other purposes such as finding the orientation of a human face in the image.

[0003] Methods are known in the art for finding human eyes in a digital image. For example, U.S. Patent 6,072,892 discloses the use of a thresholding method to detect the position of human eyes in a digital image. In this method, a scanning window scans across the entire image using a raster scanning method. A histogram extractor extracts an intensity histogram from the window as it scans across the image. Each intensity histogram is examined by a peak detector to find three peaks in the histogram representing the skin, the white of the eye, and the black of the pupil. A histogram having the three peaks identifies a location in an image that potentially defines an eye position. Eye position is determined from among the potential locations by calculating the area under the histogram associated with each potential location and by selecting the location that is associated with the histogram with the largest area.

[0004] One of the problems with this approach is that the entire image must be scanned on a pixel-by-pixel basis. Thus, a search window must be positioned at each pixel in the image and a histogram must be assembled at each pixel location. Further, the area under each histogram must be calculated and stored. It will be appreciated that this method consumes enormous amounts of computing power and reduces the rate at which images can be processed. This method can also produce a high rate of false positives.

[0005] Methods are also known to detect human eyes that have abnormally high red content. Such abnormally high red content is commonly associated with a photographic phenomenon known as red eye. Red eye is typically caused by a flash of light that is reflected by a pupil. As is described in commonly assigned and co-pending U.S. Patent application No. 08/919,560, it is known to search in images for pixels having the high red content that is indicative of red eye. Similarly, commonly assigned U.S. Patent 5,432,863 describes a user interactive method for detecting pixels in an image that have color characteristic of red eye. It will be recognized that these methods detect eyes only where red eye is present.

[0006] Thus, there is a need for a method to locate human eyes in a digital image with greater accuracy and efficiency.

[0007] The need is met according to the present invention by a digital image processing method for detecting human eyes in a digital image. This method comprises the steps of: detecting iris pixels in the image; clustering the iris pixels, and selecting at least one of the following methods to identify eye positions: applying geometric reasoning to detect eye positions using the iris pixel clusters; applying a summation of squared difference method using the iris pixel clusters to detect eye positions; and applying a summation of squared difference method to detect eye positions from the pixels in the image; wherein, the method applied is selected on the basis of the number of iris pixel clusters.

Figure 1 is a schematic diagram of an image processing system useful in practicing the present invention;

Figure 2 is a flowchart illustrating the eye detection method of the present invention;

Figure 3 is an illustration of the relationship between certain geometric parameters and an oval shaped skin color region in an image;

Figure 4 is an illustration showing, the conditional probability that a given pixel is an iris pixel stated as a function of a specific red intensity and the conditional probability that a given pixel is a non-iris pixel as a function of a specific red intensity I.

Figure 5 is a flowchart presenting the process of developing a statistical model representing the conditional probability that a given pixel is an iris pixel as a function of a specific red intensity level and a statistical model representing the conditional probability that a given pixel is a non-iris pixel as a function of a specific red intensity level;

Figure 6 is an illustration showing the iris color pixel clusters;

Figure 7 is a flowchart illustrating the process of applying the method of summation of the squared difference to eye position detection using iris pixel clusters;

Figure 8 shows an eye template and a search window centered at the center of an iris pixel cluster; and

Figure 9 shows an eye template and an image as used in the process of applying the method of summation of the squared difference to eye position detection using image pixels.

[0008] Fig. 1, shows an image processing system useful in practicing the present invention including a color digital image source 10, such as a film scanner, digital camera, or digital image storage device such as a compact disk drive with a Picture CD. The digital image from the digital image source 10 is provided to an image processor 12, such as a programmed personal computer, or digital image processing work station such as a Sun Sparc 20 workstation. The image processor 12 may be connected to a CRT display 14, and an operator interface such as a keyboard 16 and a mouse 18. The image processor 12 is also connected to a computer readable storage medium 17. The image processor

12 transmits processed digital images to an output device 19. Output device 19 can comprise a hard copy printer, a long term image storage device, a connection to another processor, or an image telecommunication device connected, for example, to the internet.

[0009] Turning now to Fig. 2, the method of the present invention will be described in greater detail. Fig. 2 is a flow chart illustrating one embodiment of the eye detection method of the present invention. Fig. 2 shows that a digital color image to be processed is first input into an iris color pixel detection step 200. In the embodiment shown, iris color pixel detection is accomplished by first detecting skin color regions in the image and then identifying iris color pixels by measuring red intensity levels from within the skin color regions.

[0010] The first step in skin color detection is color histogram equalization shown in Fig. 2 as step 201. Color histogram equalization step 201 receives images to be processed and ensures that the images are in a form that will permit skin color detection. This step is made necessary because human skin may take on any number of colors in an image because of lighting conditions, flash settings and film characteristics. This makes it difficult to automatically detect skin in such images. In Color Histogram Equalization step 201, a statistical analysis of each image is performed. If the statistical analysis suggests that the image may contain regions of skin that have had their appearance modified by lighting conditions, then such images are modified so that skin colored regions can be detected.

[0011] After the Color Histogram Equalization step, the image is searched for skin color regions in Skin Color Detection step 202. While it is possible to detect skin in a digital image in a number of ways, a preferred method for detecting skin in a digital image is to separate skin color pixels from other pixels in an image by defining a working color space that contains a range of possible skin colors collected from a large, well-balanced population of images. A pixel is then identified as skin color pixel if the pixel has a color that is within the working color space.

[0012] Skin Color Detection step 202 identifies a region of skin color pixels in the image. This region can be defined in any number of ways. In one embodiment, the skin color region is defined by a set of pixel locations identifying the pixels in the image having skin colors. In another embodiment, a modified image is generated that contains only skin color pixels. In yet another embodiment, Skin Color Detection step 202 defines boundaries that confine the skin color region in the image. It will be recognized that more than one skin color region can be identified in the image.

[0013] Oval Region Extraction step 204 examines the skin color regions detected by the Skin Color Detection step 202 to locate skin color regions that may be indicative of a face. Because the human face has a roughly oval shape, the skin color regions are examined to locate an oval shaped skin color region. When an oval

shaped skin color region is found, the Oval Region Extraction step 204 measures the geometric properties of the oval shaped skin color region. The Oval Region Extraction step 204 uses these measurements to define parameters that describe the size of the face and the location of the face within the image.

[0014] Fig. 3 is an illustration of the relationship between the geometric parameters used to define an oval shaped skin color region in the image. As is shown in Fig. 3, these parameters include Oval_top 300, Oval_bottom 302, Oval_left 304, Oval_right 306, Oval_center_row 308, and Oval_center_column 310. These parameters are used in later steps of the present method increase the efficiency of the eye detection process. It will be recognized that the method of the present invention can be practiced using skin color detection regions that have shapes that are other than oval and that other geometric parameters can be defined in association with such shapes. It will also be recognized that it is not necessary to detect an oval or other shaped area in the image. In such a case, the skin color region is examined to detect iris color pixels. Further in this case, still other parameters describing the skin color region are defined for use in the eye detection process.

[0015] After oval region extraction has been performed, the oval shaped skin color region is searched for iris color pixels. This step is performed by Iris Color Pixel Detection step 206. It will be recognized that limiting the search for iris color pixels to those pixels within the oval shaped skin color region increases the efficiency of the iris color pixel detection. It will also be recognized that there are many ways for iris pixel detection step 200 to detect pixels that are associated with an iris. Such pixels can be identified by simple color thresholding methods, model mapping and other methods well known in the art.

[0016] In one preferred embodiment, iris pixels are detected using the method claimed and described in co-pending and commonly assigned U.S. Patent application Serial No. 09/866,183 entitled Digital Image Processing Method and Computer Program Product for Detecting Human Irises in an Image. In this embodiment Iris Color Pixel Detection step 206 determines whether a pixel is an iris by measuring the red intensity of the pixel. This is done because it has been observed that a human iris has a low red intensity as compared to human skin which has a relatively high red intensity. However, in this embodiment, iris color pixels are not separated from skin color pixels on the basis of a simple threshold method. Instead, the red intensities of the pixels in the oval shaped skin colored region are used to determine the probability that each pixel is an iris and to determine the probability that each pixel is not an iris. The relationship between the probability that the pixel is an iris and the probability that the pixel is not an iris is then analyzed to determine whether the pixel is an iris pixel.

[0017] The probability that a pixel having a given red

intensity is an iris pixel is determined based upon an iris statistical model. Similarly, a non-iris statistical model is used to define the probability that a given pixel is not an iris pixel based upon the red intensity level of the pixel. The relationship between these models is non-linear as is shown by way of example in Fig. 4, which is an illustration showing, an example of a statistical model representing the conditional probability 402 that a given pixel is an iris pixel as a function of a specific red intensity and an example of a statistical model representing the conditional probability 404 that a given pixel is a non-iris pixel as a function of a specific red intensity I.

[0018] The probability analysis can take many forms. For example, the probabilities can be combined in various ways with a pixel being classified as an iris or not on the basis of the relationship between these probabilities. However, in a preferred embodiment, a mathematical construct known as a Bayes model is used to combine the probabilities to produce the conditional probability that a pixel having a given red intensity belongs to an iris.

[0019] In this embodiment, the Bayes model is applied as follows:

$$P(iris | I) = \frac{P(I | iris)P(iris)}{P(I | iris)P(iris) + P(I | noniris)P(noniris)}$$

where $P(iris | I)$ is the conditional probability that a given pixel intensity belongs to an iris; $P(I | iris)$ is the conditional probability that a given iris pixel has a specific intensity I; $P(iris)$ is the probability of the occurrence of an iris in the face oval region; $P(I | noniris)$ is the conditional probability that a given non-iris pixel has a specific intensity I; and $P(noniris)$ is the probability of the occurrence of a non-iris pixel in the face oval region. The Bayes model further applies the probability of the occurrence of an iris in a face oval region and the probability of the occurrence of a non-iris pixel in the face oval region. Using a probability analysis based on the Bayes model, a pixel is classified as an iris if the conditional probability that a pixel having a given red intensity belongs to an iris is greater than, for example, 0.05.

[0020] In the embodiment described above, only those pixels in the oval shaped skin color region defined by Oval_top 300, Oval_bottom 302, Oval_left 304, and Oval_right 306 are examined. Confining the pixels to be examined to those in the Oval shaped skin color region reduces number of pixels to be examined and decreases the likelihood that pixels that are not irises will be classified as such. It will be understood that confining the pixels to be examined to those in the skin color region, as is described in other embodiments of the present invention, conveys similar advantages. It will also be understood that it is not necessary to detect skin color regions and that the present invention can be practiced by measuring the red intensity of each pixel in the image and determining whether a pixel is an iris based upon the probability analysis described above.

[0021] Fig. 5 shows a flow chart illustrating the Iris Color Bayes Model Training step 226 used to define the statistical model used to determine whether the pixel is an iris pixel and the statistical model used to determine whether the pixel is a non-iris pixel. The method of step 226 is performed before the method for detecting irises of the present invention is used to detect irises. As is shown, in Fig. 5, a large sample of frontal face images is collected and examined. All iris pixels and non-iris pixels in the face region are then manually identified 502 and 504. Next, the conditional probability that a given iris pixel has a specific red intensity I, $P(I | iris)$ is computed and the probability of the occurrence of an iris in the face oval region, $P(iris)$ 506 is computed; then the conditional probability that a given non-iris pixel has a specific red intensity I, $P(I | noniris)$ is computed and finally the probability of the occurrence of a non-iris pixel in the face oval region, $P(noniris)$ 508 is computed. The computed statistical models of iris and non-iris are used in the Bayes model to produce the conditional probability that a given pixel intensity belongs to an iris $P(iris | I)$ 510. In certain embodiments the Bayes model can be used to generate a look-up table to be used in Iris Color Pixel Detection step 206.

[0022] After the Iris Color Pixel Detection step 206 identifies the location of the iris pixels in the image, the iris color pixels are then assigned to clusters. This is done by Iris Pixel Clustering step 208. A cluster is a non-empty set of iris color pixels with the property that any pixel within the cluster is also within a predefined distance to another pixel in the cluster. One example of a predefined distance is one thirtieth of the digital image height. The Iris Pixel Clustering step 208 of Fig. 2 groups iris color pixels into clusters based upon this definition of a cluster. However, it will be understood that pixels may be clustered on the basis of other criteria.

[0023] Under certain circumstances, the definition of a cluster of iris color pixels may be broad enough to include clusters that are invalid. In such a circumstance, as is shown in Fig. 2, the step of validating the clusters is included as step 209. A cluster may be invalid because, for example, it contains too many iris color pixels or because the geometric relationship of the pixels in the cluster suggests that the cluster is not indicative of an iris. For example, if the ratio of a cluster's height to its width is determined, and if the ratio is greater than two, then this cluster is invalid. Invalid iris pixel clusters are removed from further consideration. Accordingly, in the portions of the description that follow, valid iris pixel clusters will be referred to simply as iris pixel clusters.

[0024] The number of iris pixel clusters "n" is counted in step 210. The number of iris pixel clusters "n" is used by decision step 210 to select from among two paths for detecting eyes in an image. If the number of iris pixels "n" is less than two then the process is branched to step 224 which is to be described later. If the number of iris color pixels "n" is at least two, then the process branches to step 212 to find the centers of the clusters. The center

of a cluster is determined as the center of mass of the cluster. The center position of the clusters is calculated with respect to the origin of the image coordinate system. For the purposes of this measurement, the origin of the image coordinate system is at the upper left corner of the image boundary.

[0025] After the center of each of the iris pixel clusters is located, geometric reasoning is applied to attempt to detect eyes based on the geometric relationship between the iris pixel clusters. As is shown in Fig. 6, if there are only two clusters remaining, with one in the left-half 604 and one in the right-half 606 and if the horizontal distance between the centers of the two clusters is less than 0.4 times the distance between Oval_right 306 and Oval_left 304 and if the vertical distance between the centers of the two clusters is less than one tenth of the distance between Oval_top 300 and Oval_bottom 302, then the center positions of these two clusters are treated as the eye positions.

[0026] It will be appreciated that this analysis can be completed very rapidly. If this analysis is successful, then no further action is required to detect eye positions. Accordingly, a detection result check is done in step 216 to see if eye positions are detected. If eye positions are detected, then the eye detection process stops. If no eye positions are detected, then the process goes to step 218.

[0027] In step 218 the summation of squared difference method is used to search the image for eye positions. In general, the summation of the squared difference method involves calculating the summation of the squared difference of the intensity values of the corresponding pixels in an eye template and a patch of the image that has the same size as the template. In this method, each pixel in the patch of pixels has a corresponding pixel in the template. The difference between the intensity level of each of the corresponding pixels is calculated. Each difference is then squared. The sum of each of the squared differences for each of the pixels in the set is then calculated. This summation of the squared differences provides a relative measure of the degree of correspondence between each of the pixel sets measured and the template. If no eye positions are detected 220, then the process goes to 224.

[0028] As applied in the present invention, summation of the squared difference values are calculated for each pixel in each window in each half-region. These values are compared and the cluster having the lowest relative summation of the squared difference value is selected identified as an eye location for the respective half region. This process is performed separately on the clusters of the left and the right-half regions of the image in the manner described below.

[0029] It will be noted that while the present invention has been described as using the summation of the squared difference method to identify the best relative correlation between the average eye template and each of the patches of pixels, other methods including the

mean-squared-error method can be used for this purpose.

[0030] In one embodiment of the method of the present invention, the parameters calculated in the Oval Region Detection Step 204 are used to increase the efficiency of the use of the summation of the squared difference method by reducing the number of locations in the image at which the summation of the squared difference must be calculated. As is shown in Fig. 6, in this embodiment, the Oval_Center_Column 310 is used to divide the oval region into a left-half region 604 and a right-half region 606. As is also shown in Fig. 6, iris pixel clusters 600 and the center position 602 of the iris pixel clusters 600 are positioned in either the left-half or right-half regions 604 and 606 separated by the Oval_Center_Column 310.

[0031] Step 218 conducts a left-eye position search using the summation of squared difference method and the pixel clusters 600 located on the left-half region 604. Step 218 also conducts a right-eye position search in the right-half region using the summation of squared difference method and the iris pixel clusters 600 located in the right-half region 606.

[0032] Referring now to Figs. 7 and 8, the process for selecting one cluster from the clusters in a half region will be described. The eye position search process is started by centering 70 a window 800 at the center of each cluster 802 in a respective half-region. The default size for window 800 is one twentieth of the size of the image 804. The operation of calculating the summation of the squared differences 72 is then performed on each of the pixels in each window 800. The position of the pixel having the lowest summation of squared difference value in each window 800 is recorded 76. When a summation of the squared difference value has been calculated for every pixel in every window of the half-region 78, the position of the pixel having the lowest summation of squared difference value is recorded 79. This is the estimated eye position for a half-region. This process is repeated for the remaining half-region. If two eyes are detected by this process, then the method is ended.

[0033] It will be appreciated that the summation of the squared difference method of step 218 can also be performed without the use of Oval Shaped Skin Color Extraction. In such an embodiment, the skin color region can be divided into a left-half region and a right-half region. Iris pixel clusters can then be divided into left-half region and right-half region clusters. The summation of the squared difference method can then be applied as described above.

[0034] However, if eyes have not been detected 220 after the execution of step 218, or if less than two iris pixel clusters are detected at step 209 then the eye detection process goes to step 224. Step 224 operates in a manner that is similar to step 218. However, as is shown in Fig. 9, the entire image 900 is divided and a summation of the squared difference is calculated for every pixel of the image 904 in the left-half 908 and right-

half 910 region respectively.

[0035] It will be understood that other methods can be used in place of the summation of squared difference method to determine the relative correlation between a patch of an image and an eye template. One example is the mean-squared error method. This method is well known in the art.

[0036] It will be appreciated that the present invention provides three distinct steps methods to detect eyes in an image; geometric reasoning 212 and 214, summation of the squared difference using iris pixel clusters 218 and summation of the squared difference using image pixels 224. It will also be understood that geometric reasoning is the simplest and most efficient of these methods. This is because geometric reasoning provides the most efficient processing method and because geometric reasoning is applied only to iris pixel clusters. These clusters are relatively small in number when compared to the number of pixels in the image.

[0037] In contrast, applying a summation of the squared difference method to each of the pixels in an image as is required in step 224 is a computationally heavy step, requiring many processing steps and calculations to determine whether a single pixel in the image is an eye position. Further, the method of step 224 must be applied to all of the non-iris pixels in an image. In a currently common format, digital images are being acquired by 2.1 megapixel cameras. Further, cameras having as many as 16 megapixels have been demonstrated. Thus, it is clear that using step 224 to detect eye positions in an image will require, literally, hundreds of millions of operations to process a single image. This is a time and computer intensive process.

[0038] As an intermediate approach, step 218 applies a computationally heavy summation of the squared difference method, but limits the application of this method to the pixels in the windows defined about the iris pixel clusters. This substantially reduces the number of pixels to which the summation of the squared difference method must be applied and therefore makes the application of the summation of the squared difference method 220 less computer intensive than the summation of the squared difference method of step 224.

[0039] It will also be understood that the method of the present invention provides a way for automatically selecting between these eye detection methods and combining these methods in a manner that uses the number of iris pixel clusters to select the most efficient method for detecting eyes in the image.

[0040] The subject matter of the present invention relates to digital image understanding technology, which is understood to mean technology that digitally processes a digital image to recognize and thereby assign useful meaning to human understandable objects, attributes or conditions and then to utilize the results obtained in the further processing of the digital image.

[0041] Thus, the present method and computer program product meets the objective of this invention.

Claims

1. A digital image processing method for detecting human eyes in a digital image, comprising the steps of:

detecting iris pixels (200);
clustering the iris pixels (208);
determining the number of iris pixel clusters (210);
selecting at least one of the following methods to identify eye positions in an image (210):

- i) applying geometric reasoning to detect eye positions using the iris pixel clusters (212, 214);
- ii) applying a summation of squared difference to detect eye positions based the iris color pixel clusters (218);
- iii) applying a summation of squared difference method to detect eye positions from the pixels in the image (224);

wherein the applying step is selected on the basis of the number of iris pixel clusters.

2. The method of claim 1, wherein applying step iii) is selected when less than two iris pixel clusters are detected.

3. The method of claim 1, wherein applying step i) is selected when at least two iris pixel clusters are detected.

4. The method of claim 3, wherein method i) is first applied and method ii) is subsequently applied in the event that method i) does not detect at least two eye positions.

5. The method of claim 4, wherein method ii) is applied and method iii) is subsequently applied in the event that method ii) does not detect at least two eye positions.

6. A digital image processing method for detecting human eyes comprising the steps of:

detecting iris pixels;
clustering iris pixels;
determining the number of iris pixel clusters;
selecting a method for detecting eye positions based upon the number of iris pixel clusters;
using the selected method to detect eye positions.

7. The digital image processing method of claim 6, wherein the step of selecting a method for detecting eye positions based upon the number of iris pixel clusters comprises selecting a series of eye detec-

tion steps.

8. The digital image processing method of claim 6,
wherein the step of selecting a method for detecting
eye positions is based upon the number of iris pixel
clusters and the time required to execute each
method. 5
9. The digital image processing method of claim 6,
wherein the step of selecting a method for detecting
eye positions is based upon the number of iris pixel
clusters and the computational efficiency of the
method selected. 10
10. The digital image processing method of claim 6, 15
wherein the step of selecting a method for detecting
eye positions is based upon the number of iris pixel
clusters and the probability that each method will
detect eyes.

20

25

30

35

40

45

50

55

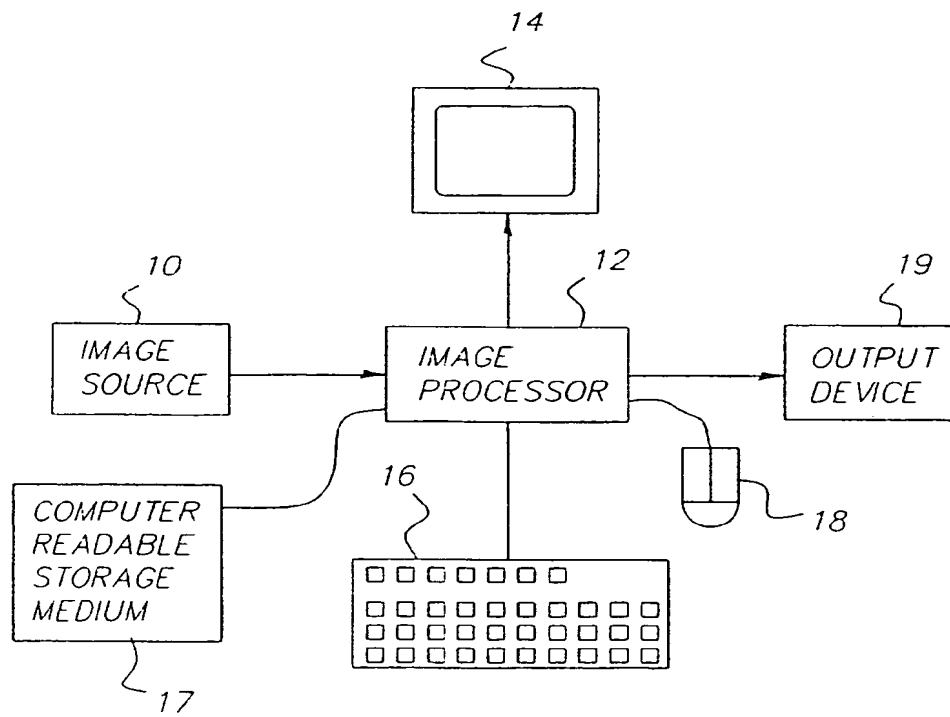


FIG. 1

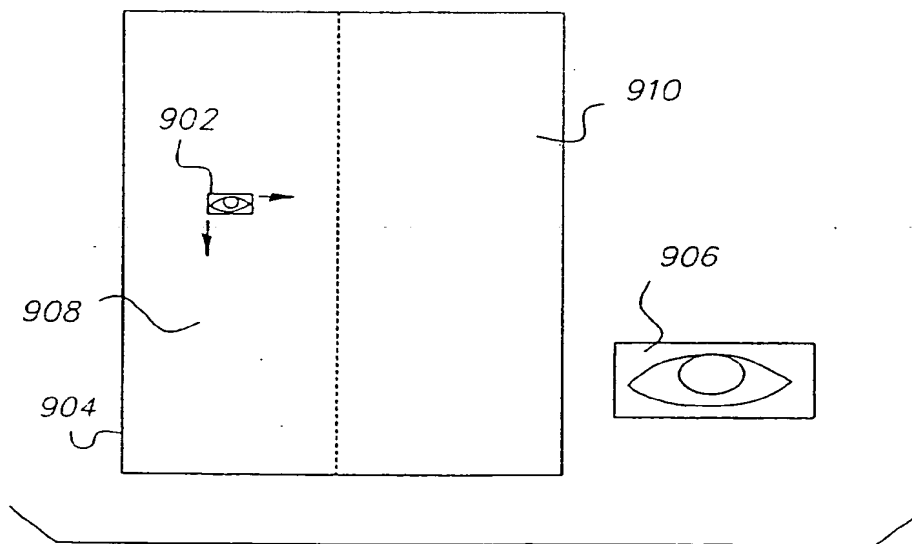
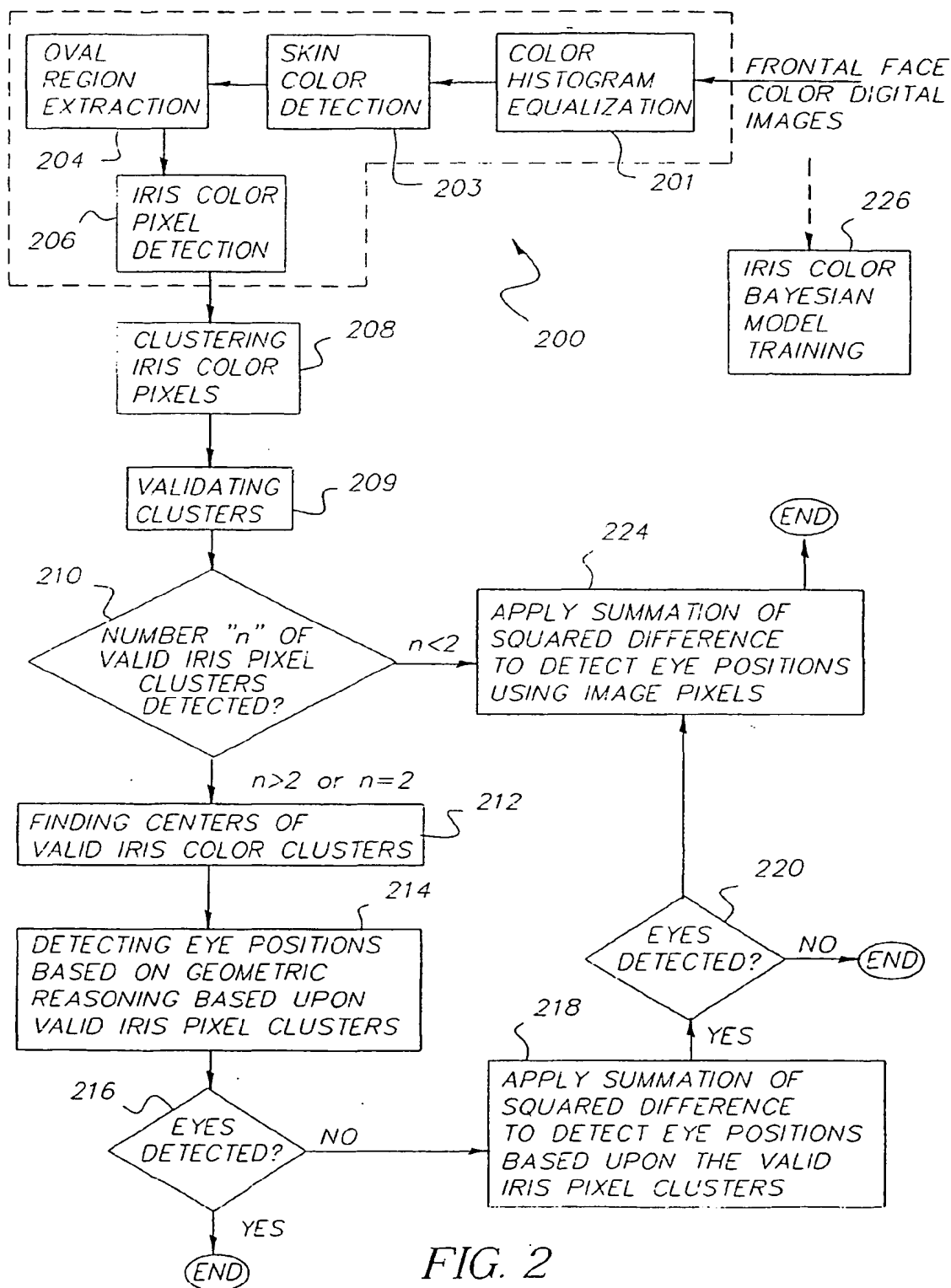


FIG. 9



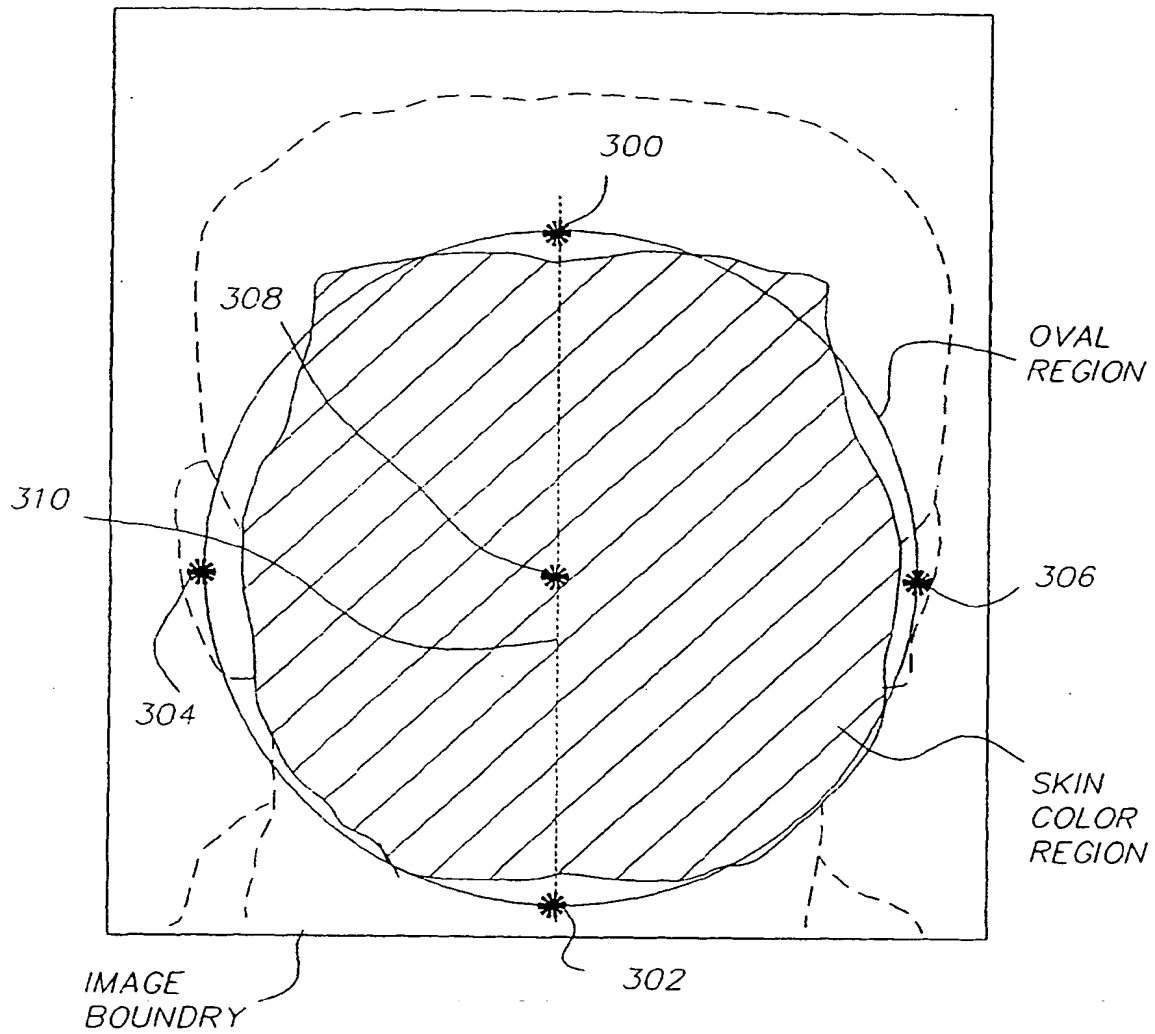


FIG. 3

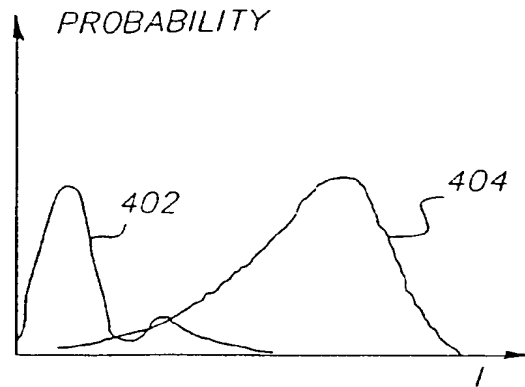


FIG. 4

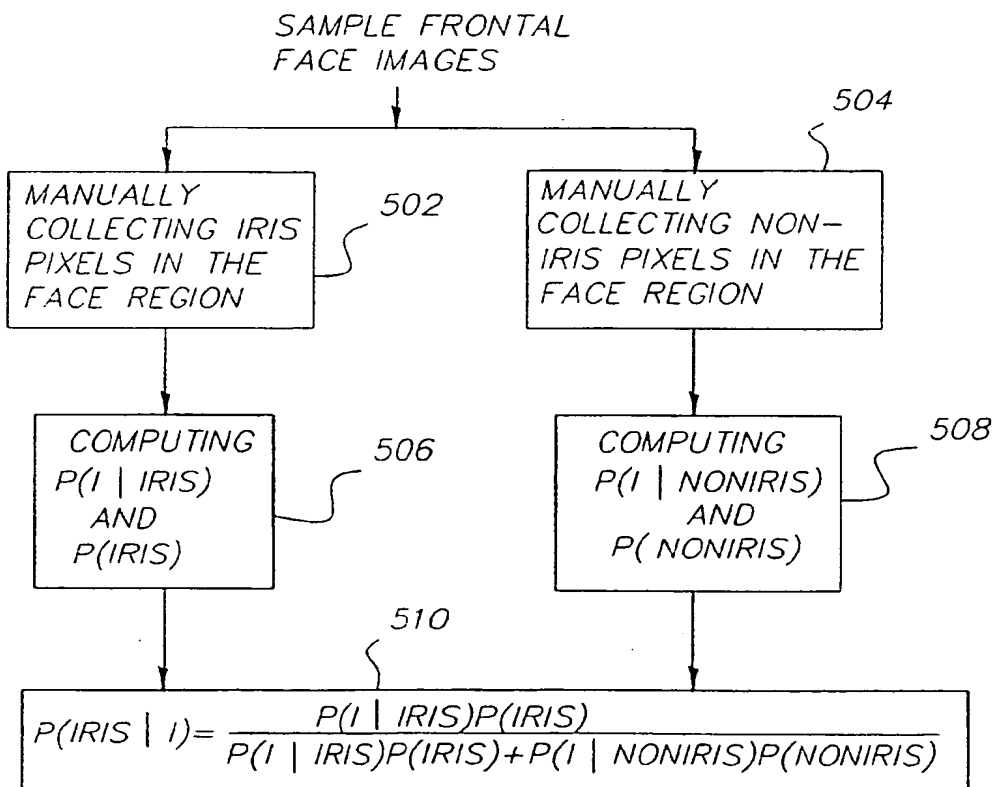


FIG. 5

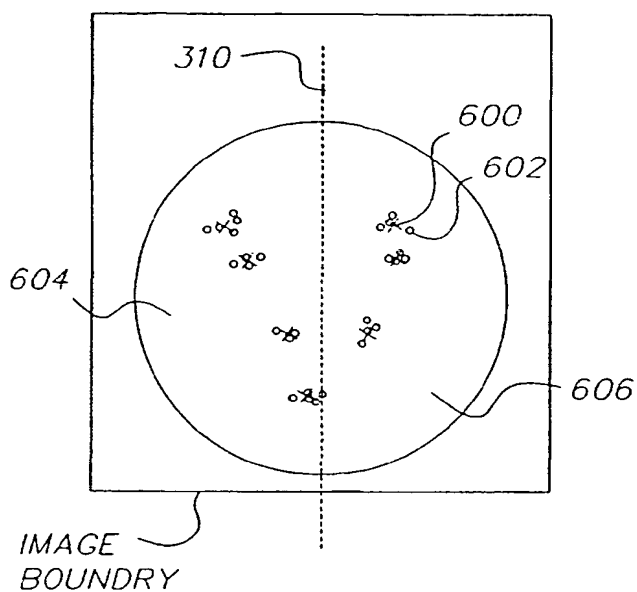


FIG. 6

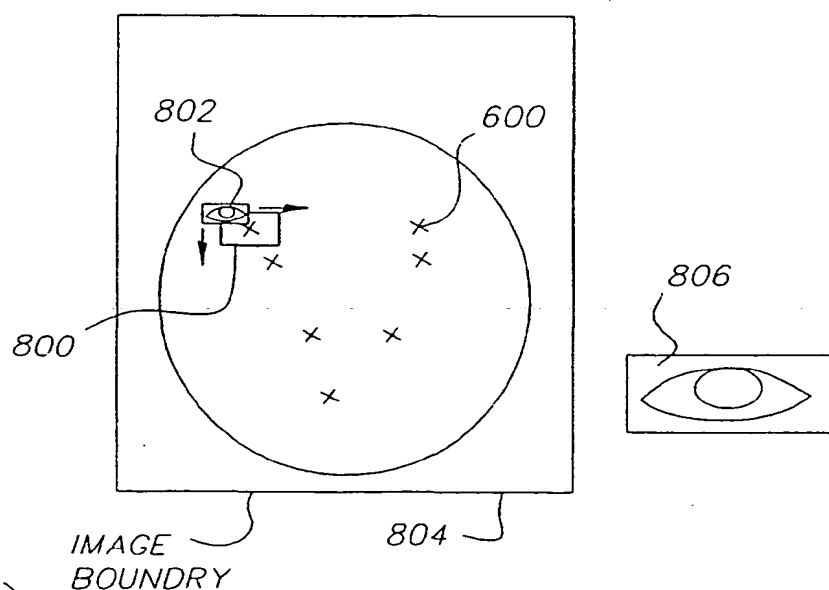
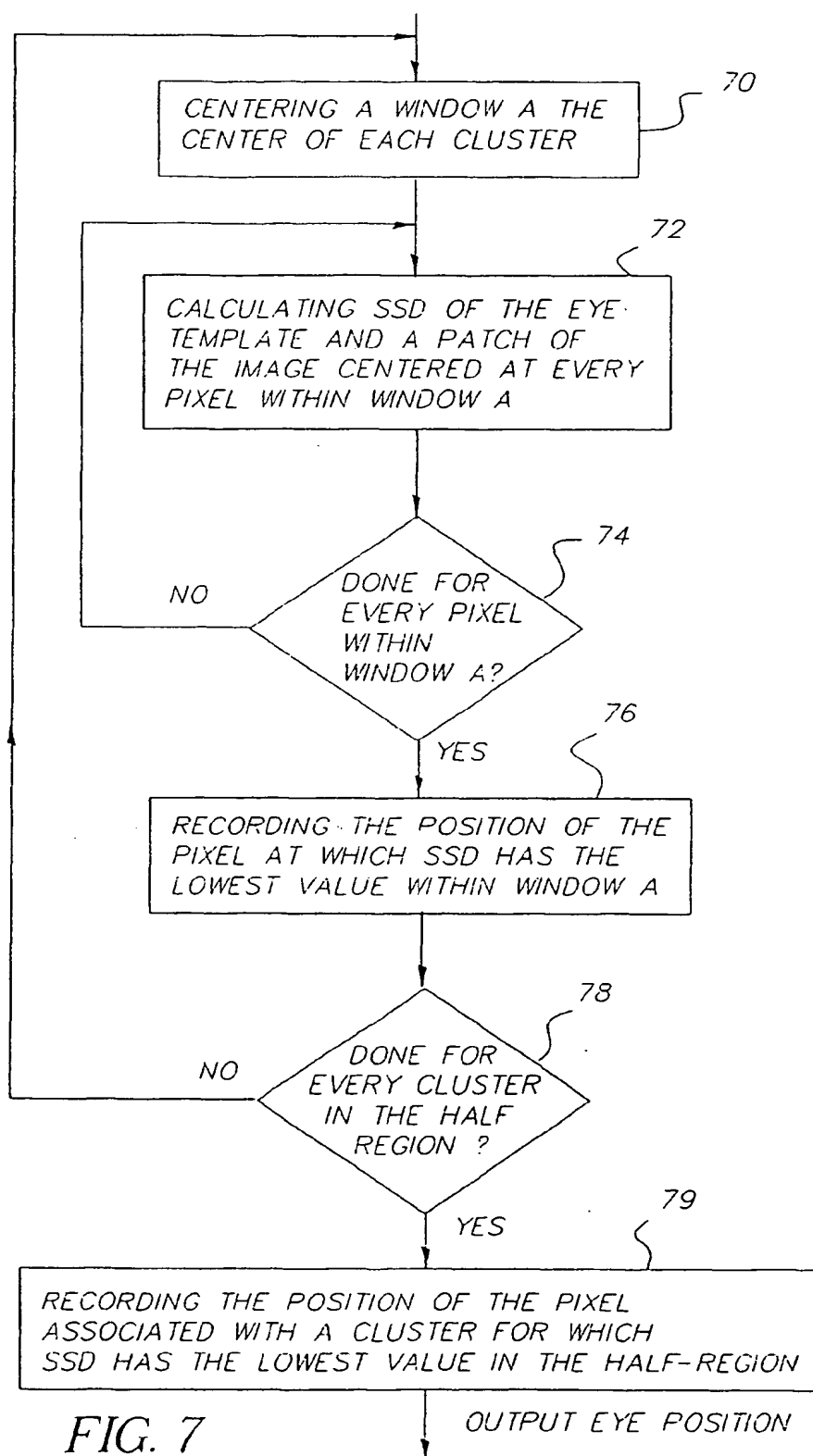


FIG. 8



THIS PAGE BLANK (USPTO)